Am J Cancer Res 2021;11(6):2684-2695
www.ajcr.us /ISSN:2156-6976/ajcr0129815

Original Article

IL-17 signaling pathway plays a key role in laryngeal squamous cell carcinoma with ethnic specificity

Li Qi[1], Wenzhao Bao[2], Wei Li[1], Xiaoxu Ding[1], Aihui Yan[1]

[1]*Department of Otorhinolaryngology, The First Hospital of China Medical University, Shenyang 110001, Liaoning Province, China;* [2]*Department of Anesthesiology, Affiliated Hospital of Inner Mongolia University for The Nationalities, Inner Mongolia, China*

Received January 9, 2021; Accepted April 25, 2021; Epub June 15, 2021; Published June 30, 2021

Abstract: Laryngeal squamous cell carcinoma (LSCC) is a common aggressive head and neck squamous cell carcinoma (HNSCC) and racial disparities have been reported to exist in it. However, its molecular mechanism and associated ethnic specificity are still unclear. Here, we leveraged mRNA expression data from 2 gene expression omnibus datasets (GSE142083 & GSE117005) of Chinese samples and the cancer genome atlas (TCGA) datasets of Caucasian samples to demonstrate the expression signature of LSCC. The GSE142083 dataset was used as the discovery set since it had 53 pairs of LSCC tissues and matched adjacent normal tissues, and the GSE117005 dataset was treated as the validation set with 5 pairs of tissues. Differential gene expression analysis and enrichment pathway analysis were performed. Besides, we employed weighted gene co-expression network analysis to identify hub genes in validated pathways. The TCGA datasets were used to evaluate ethnic specificity. Immunohistochemistry (IHC) was employed to further validate the hub gene. Overall, the IL-17 signaling pathway was significantly enriched for upregulated genes in two Chinese datasets while not in TCGA datasets; and IL17RC, MAPK3, S100A8, MMP3, CXCL8, and TNFA1P3 were hub genes regulating such pathway. Therein, IL17RC might be the most important one and the IHC results displayed that the IL17RC gene upregulated in the LSCC tissue. IL-17 signaling pathway has an ethnicity-specific effect in LSCC where it is upregulated in the Chinese while not in the Caucasians and IL17RC might play a key role. Targeting genes located in the IL-17 signaling pathway may be beneficial for Chinese LSCC patients.

Keywords: Laryngeal squamous cell carcinoma, gene expression, IL-17 signaling pathways, ethnic specificity

Introduction

Laryngeal squamous cell carcinoma (LSCC) is one of the most common head and neck tumors, and racial disparities have been observed between African Americans and Caucasians [1]. Patients with LSCC usually have difficulty in vocalization, breathing, and swallowing. Surgery, together with radiation and systemic therapy, has been the mainstay for LSCC but the overall survival rate is still far from satisfaction [2]. Current studies indicated that infection-free LSCC tended to be burdened with tobacco-associated mutations in TP53 and CDKN2A [3] and overexpression of SKA3 was confirmed to be associated with tumor proliferation and chemoresistance in LSCC patients [4]. However, no ideal biotherapy has been approved for LSCC and its underlying molecular mechanism still needs further exploration.

The human immune system has been reported to be closely associated with the initiation and progression of cancer, and cancer often occurs in those with compromised immunity [5]. Meanwhile, many drugs that aim to modulate anti-tumor immune response have been developed. Although immunotherapeutic drugs such as nivolumab and pembrolizumab have been approved for squamous cell carcinoma of the head and neck (HNSCC) [6], the immunotherapy for LSCC is still under exploration and its hidden molecular mechanism needs further investigation. As a substantial component of the inflammatory and immune system, the IL-17 signaling pathway has been reported to be associated with the initiation and progression of cancer [7]. For example, overexpression of IL-17 from gamma delta T cells and neutrophils conspired to promote breast cancer metastasis [8]. Also, IL-17 upregulates PD-L1 expression in prostate cancer [9], and targeting the Th17-

IL-17 pathway may help in the prevention of micro-invasive prostate cancer [10].

However, whether the IL-17 signaling pathway involves in LSCC is still unknown and only one study demonstrated the IL-17 signaling pathway was activated in idiopathic subglottic stenosis [11]. Another study indicated IL-17 was overexpressed in LSCC patients compared to those with polypus of the vocal cord [12], but no systematic study was conducted and existing evidence is far from enough. With the accumulation of biodata and various bioinformatics methods, it is possible to perform comprehensive data mining to reveal the molecular mechanism for cancer and find possible therapeutic targets. Here, we integrated gene expression data from gene expression omnibus (GEO) and the cancer genome atlas (TCGA) to unveil potential molecular mechanisms, especially those associated with immunity; and further to discover hub genes, hoping to find potential targets. Furthermore, we also explored the ethnicity-specific effect of the IL-17 signaling pathway and discussed the potential racial disparities in LSCC patients.

Methods

Data source & main design

Two latest GEO datasets of Chinese samples (GSE142083 [10] & GSE117005 [13]) were obtained. The former dataset consisted of 53 pairs of LSCC tissues and matched adjacent non-LSCC tissues while the latter contained 5 pairs of matched tissues. TCGA datasets with the larynx as the primary site were obtained. After removing those without matched adjacent normal tissues and non-Caucasian samples, 11 pairs of matched tissues were kept for further analysis. All datasets were generated from the RNA-Seq experiment.

All raw datasets were downloaded in the FPKM (Fragments Per Kilobase per Million) format and they were transformed to TPM (Transcripts Per Million). All probes were annotated with gene symbols and probes without annotation were excluded. Besides, only the probe with the highest expression level was reserved if different probes were annotated to the same gene. The annotation file was downloaded from GENCODE (https://www.gencodegenes.org/).

Considering the sample size and number of annotated probes, the discovery stage was performed in the dataset GSE142083 to find differentially expressed genes and possible pathways. Meanwhile, weighted gene co-expression network analysis (WGCNA) [14] was also conducted to bring out key modules and hub genes closely associated with LSCC in this dataset. Then, the dataset GSE117005 was utilized to validate the findings in GSE142083 considering its relatively small sample size and fewer annotated probes. The validated results were further tested in TCGA datasets to compare the ethnic specificity.

Differentially expressed gene analysis & pathway enrichment analysis

A paired design RNA-Seq experiment of LSCC tissues and matched adjacent normal tissues was performed to detect genes differentially expressed between LSCC and normal tissues, adjusting for any differences between the patients. The gene was removed if its expression level was too low, usually with a CPM (Counts Per Million) less than 1 in more than 99% samples. Generalized linear methods were adopted to determine differentially expressed genes and likelihood ratios were conducted to test LSCC versus normal tissue differences, and the *p*-value was adjusted using the false discovery rate (FDR) method. To bring out more differentially expressed genes, the cutoff of log fold change (logFC) was calculated using the formula below:

$$\text{cutoff} = \min\left(\left|\overline{logFC}\right| + 2 * SD(logFC), 1.5\right)$$

Here, the "min" represents the minimum value of the two numbers and "SD" is the standard deviation.

All differentially expressed genes were tagged by 'UP' or 'DOWN' and extracted for pathway enrichment analysis. We enriched upregulated and downregulated genes in the Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) databases separately to discover significantly changed pathways [15, 16]. Meanwhile, all differentially expressed genes were pooled to find more potential pathways. Pathways with a *p*-value adjusted by FDR less than 0.05 were defined to be significantly enriched in LSCC tissues. The differentially expressed genes were analyzed in R software

using package "edgeR" [17] and pathway enrichment analysis was performed with package "clusterProfiler" [18].

Validation for differentially expressed genes and pathways

The dataset GSE117005 was employed to validate the differentially expressed genes and enriched pathways from GSE142083. In this validation stage, we defined the pathway, which was significant in the GSE142083 dataset, to be the validated pathway if its *p*-value was less than 0.05. Similarly, differentially expressed genes with the same direction as those in the GSE142083 dataset were thought to be validated genes.

Weighted gene co-expression network analysis

The gene expression data from GSE142083 was used to perform WGCNA. Hierarchical cluster analysis was performed to detect outliers in total samples and outliers with minimal cluster size less than 2 would be removed in further analysis. Then, we picked up the best soft threshold 5 to ensure a scale-free R^2 of more than 0.9. The genes were hierarchically clustered based on their topological overlap. A weighted gene co-expression network was constructed to identify the modules closely associated with LSCC, and we further related identified modules to external sample information and appraised the association between gene significance and module membership. Here, we used automatic, one-step network construction and module detection. Hub genes could be identified in significant modules according to their intramodular connectivity. Specifically, we mainly focused on identifying key genes in the validated pathways. This analysis was performed in R software using the package "WGCNA".

Evaluation of ethnicity-specific effect in LSCC

Considering the racial disparities in LSCC, we analyzed the 11 pairs of LSCC tissues and matched adjacent normal tissues from Caucasian samples in TCGA datasets. Differentially expressed genes were detected and related pathways were enriched using the same methods abovementioned. Then, we simply compared the results from Chinese and Caucasian samples to evaluate ethnic specificity.

Tissue preparation and immunohistochemistry (IHC)

In total, 18 paired of LSCC tissue and matched adjacent normal tissue were collected by the First Affiliated Hospital of China Medical University. After dewaxing by Xylene-Xylene-100% alcohol-100% alcohol-95% alcohol-90% alcohol-80% alcohol-70% alcohol, the sections were placed in 3% hydrogen peroxide and incubated for 10 min to eliminate the activity of endogenous peroxidase. Then, the sections were placed in 10 mM citrate buffer solution (pH = 6.0) and microwaved for boiling for antigen retrieval. After cooling to room temperature, the citrate buffer was poured off; and the slides were placed in phosphate buffer solution (PBS) for 5 min. Then, serum was added immediately to close up some non-specific sites; and then placed in a 37°C incubator for half an hour. Next, the sections were incubated with an anti-IL17RC rabbit polyclonal antibody (Abcam, ab96310) at 4°C for 24 h, with the optimal dilution ratio of 1:100. After rinsing thoroughly with PBS, the sections were immunohistochemically colored with horseradish peroxidase-labeled goat anti-rabbit polymers. Finally, the positive expression of the IL17RC protein was observed with the incubated chromogenic agent, 3',3'-diaminobenzidine, and the cell membranes were counterstained with Mayer's hematoxylin. Two independent experimenters scored the immunohistochemical staining results of the sections blindly. When the experimenters scored sections differently, the sections were re-examined by a third experimenter to reach an agreement. Signal strength was scored in the following manner: no coloring or weak (0); light yellow (1); brownish yellow (2); brownish brown (3). The staining distribution score is based on the percentage of stained positive cells: 0 (<5%), 1 (5%-25%), 2 (26%-50%), 3 (>50%). The total score multiplied the results of the two scores: <3 as negative and ≥3 as positive. Fisher's precision probability test was performed to test the difference of IL17RC protein level between LSCC and normal tissues.

All patients signed an informed consent form. The study was conducted in compliance with the Declaration of Helsinki and was approved by the Ethics Review Committee of the First Hospital of China Medical University.

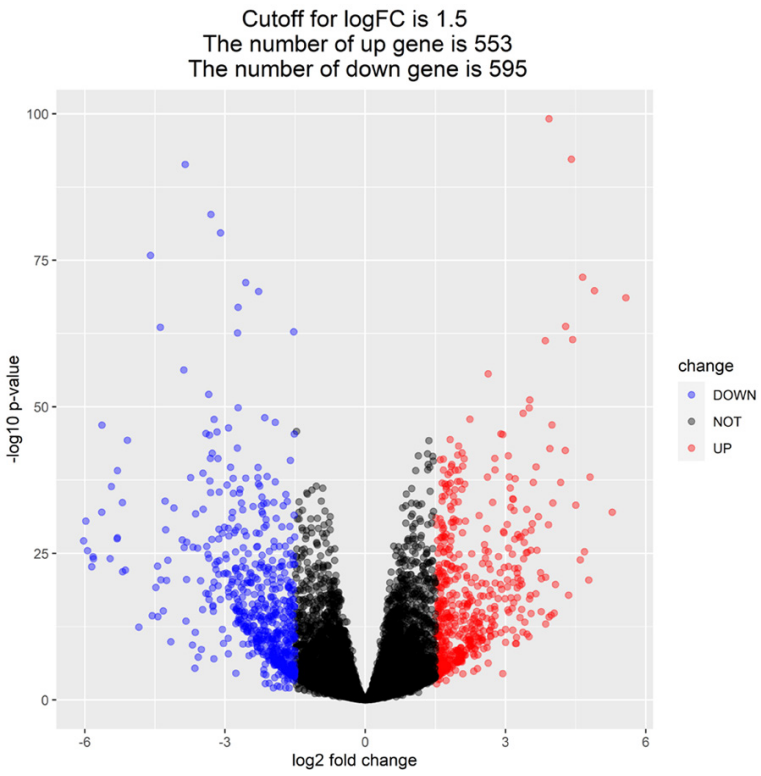

Figure 1. Volcano plot for differentially expressed genes in GSE142083.

Results

Differentially expressed genes and enriched pathways from the GSE142083 dataset

In the dataset GSE142083, 18349 mRNAs annotated with gene symbols were kept for differential gene expression analysis. The cutoff of logFC was 1.5. A total of 553 genes were upregulated and 595 genes downregulated in LSCC tissue (**Figure 1**). The top 5 upregulated genes were MMP1, SPRR2G, PAGE2, PRR9, S100A7A; and the top 5 downregulated genes were C6orf58, PIP, PRB3, PRH1, and PRH2. Therein, MMP1 and S100A7A participate in the IL17 signaling pathway. The top 50 changed genes were listed in Table S1.

When these genes were enriched in the KEGG pathway database, several upregulated pathways were significant including cytokine-cytokine receptor interaction, cell cycle, ECM-receptor interaction, IL-17 signaling pathway, protein digestion and absorption, and other pathways (**Figure 2A**); and salivary secretion pathway was downregulated in LSCC, together with cAMP signaling pathway and others (**Figure 2B**). When pooling differentially expressed genes together, cytokine-cytokine receptor interaction and IL-17 signaling pathway were the top 2 significant pathways. Therein, 14 genes were enriched to the IL-17 signaling pathway, including MMP9, DEFB4A, MMP13, MMP3, S100A7, S100A7A, MMP1, IL1B, CXCL8, CXCL10, CCL7, CXCL5, CSF2, and CCL11 (Figure S1A). In the GO database, the upregulated pathways included skin development, epidermis development, keratinocyte differentiation, epidermal cell differentiation, keratinization, and others (**Figure 2C**) while pathways associated with glucose metabolism were downregulated in LSCC, such as glycosylation and protein glycosylation (**Figure 2D**). Besides, we have provided top 50 pathway enrichment analysis results from KEGG and GO respectively (Table S2) and other visualized pathways were provided in Figure S1.

Furthermore, we put enriched pathways into a network with edges connecting overlapping gene sets to identify functional modules easily (**Figure 3**). We observed that cytokine-cytokine receptor interaction and IL-17 signaling pathway shared a relatively large portion of genes from the KEGG database, suggesting these two signaling pathways might interact with each other in LSCC (**Figure 3A**). Additionally, two biological processes, including epidermis development (**Figure 3C**) and O-glycan processing (**Figure 3D**), were possible key pathways in LSCC according to the GO database.

Validation in GSE117005 dataset

Only 5367 mRNAs were annotated with gene symbols in dataset GSE117005. Its logFC cutoff was 0.37 and 14 genes were upregulated and 27 downregulated (**Figure 4**), and 26 were significant in previously analyzed dataset GSE142083 (Table S3). Of 14 upregulated genes, 12 genes were also upregulated in the GSE142083, including ALOXE3, GRIN2D,

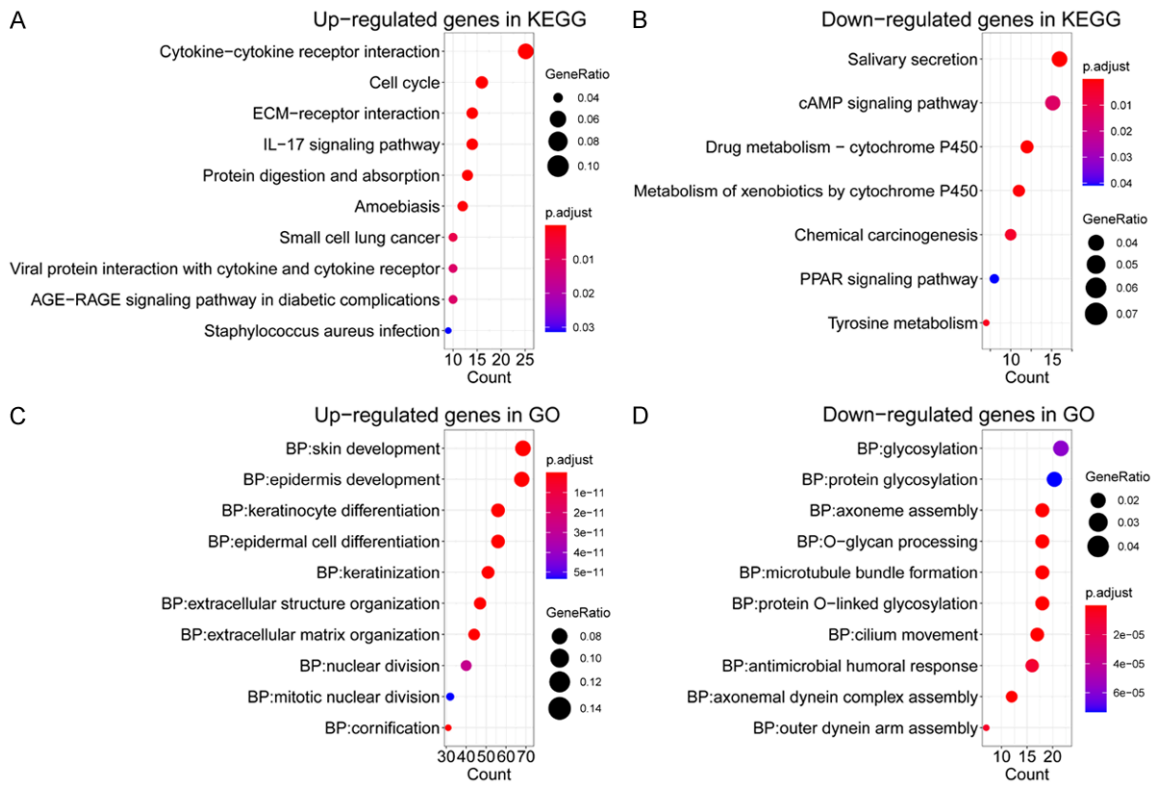

Figure 2. Pathway enrichment analysis results for differentially expressed genes in GSE142083. (A, B) are from KEGG where (A) represents upregulated genes and (B) downregulated genes; and (C, D) from GO where (C) represents upregulated genes and (D) downregulated genes.

HOXA13, HOXC10, HOXD13, LRRC15, MAGEA1, MMP1, SPRR2G, SPRR4, XCL1.

When enriching these 41 differentially expressed genes in KEGG and GO pathway databases, the results were different from those from GSE142083 with nitrogen metabolism dominating these pathways in KEGG and calcium ion channel dominating the pathways in GO; and only the IL-17 signaling pathway was still significantly upregulated (**Figure 5**). Two genes functioning in the IL-17 signaling pathway were both upregulated in two datasets, namely MMP1 and CXCL10. Intriguingly, MMP1 was the top gene in both datasets, indicating that matrix metalloproteases may play a key role in the initiation and progression of LSCC.

Weighted genes co-expression network construction and module detection

In WGCNA, 16 samples were detected to be outliers in cluster analysis and removed in further analysis (Figure S2) and the selected soft threshold was 5 with scaled free R^2 more than 0.9 (Figure S3). In the constructed co-expression network, it was clear that two modules, midnight blue and yellow, were the top 2 significant modules in the network (**Figure 6**). When relating the modules to clinical traits, the yellow module was positively associated with LSCC while the midnight blue module was negatively associated with LSCC (Figure S4). Furthermore, the yellow module membership was strongly in positive correlation with gene significance for LSCC (r^2 = 0.8, *p*-value <1×10^{-200}, **Figure 7**). In the yellow module, 103 genes were differentially expressed in previous analysis for the GSE142083 dataset (Table S4). When enriched in KEGG and GO database, cell cycle and nuclear division were the most significant ones in KEGG and GO respectively (Figure S5). Cell cycle and nuclear division have been closely linked to oncogenesis.

Then, we selected the genes located in the IL-17 signaling pathway to construct the protein-protein interaction (PPI) network; and IL17RC, MAPK3, S100A8, MMP3, CXCL8, and

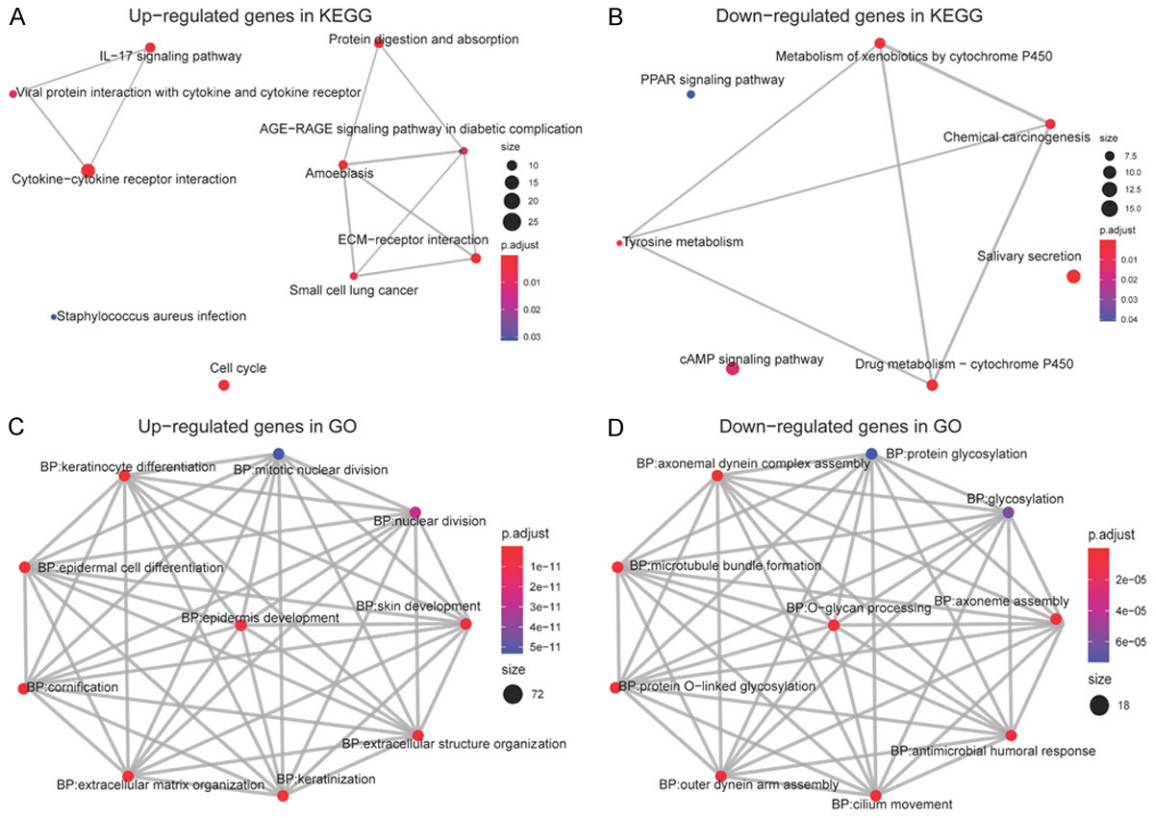

Figure 3. Enrichment map to visualize function modules in GSE142083. (A, B) are from KEGG where (A) represents upregulated genes and (B) downregulated genes; and (C, D) from GO where (C) represents upregulated genes and (D) downregulated genes. The description of dots is identical to that of **Figure 2**.

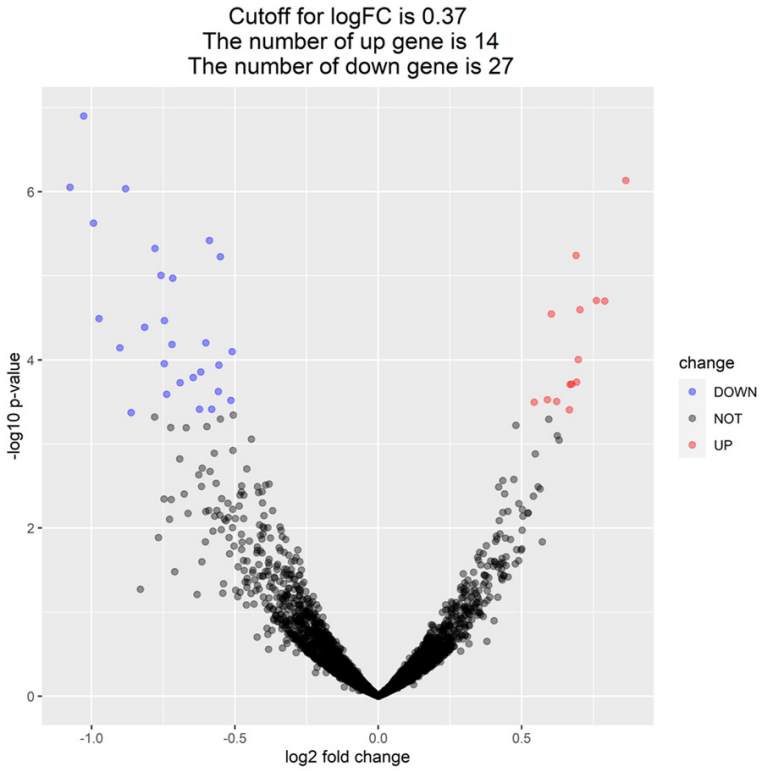

Figure 4. Volcano plot for differentially expressed genes in GSE117005.

TNFA1P3 were the hub genes in the IL-17 signaling pathway (**Figure 8**).

Ethnicity-specific effect in LSCC

Furthermore, the analytic results of Caucasian samples in TCGA datasets indicated a rather different expression pattern from Chinese samples. With 1.5 as the cutoff of logFC, 896 genes were upregulated and 926 were downregulated. Among them, 28.91% of upregulated genes and 27.32% of downregulated genes were found in GSE142083, suggesting a poten-

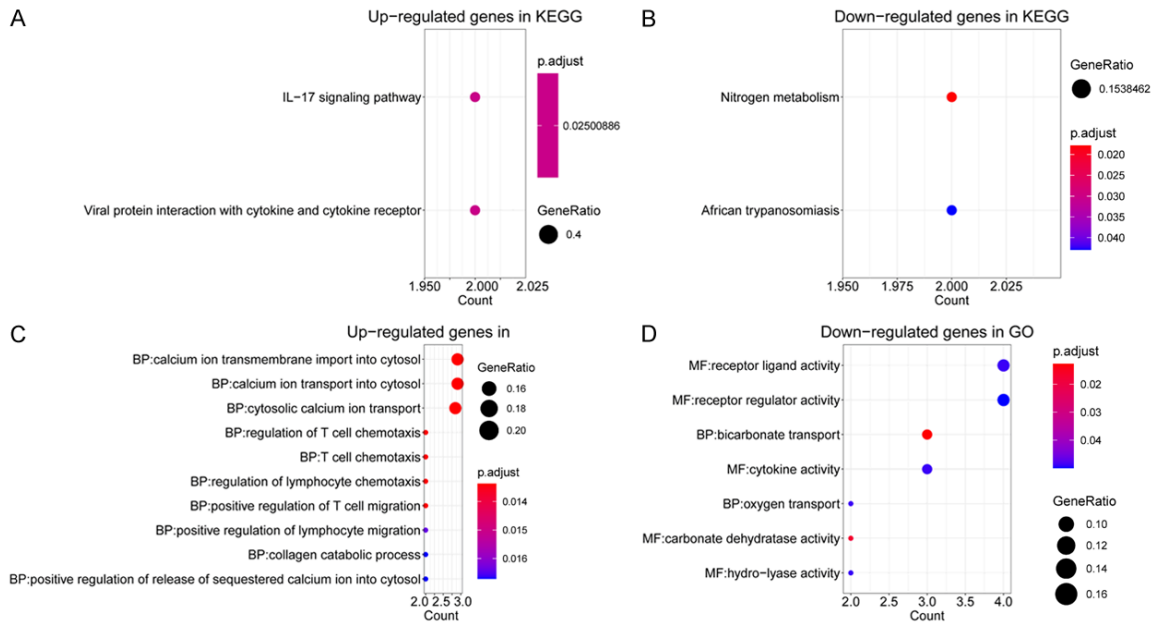

Figure 5. Pathway enrichment analysis results for differentially expressed genes in GSE117005. (A, B) are from KEGG where (A) represents upregulated genes and (B) downregulated genes; and (C, D) from GO where (C) represents upregulated genes and (D) downregulated genes.

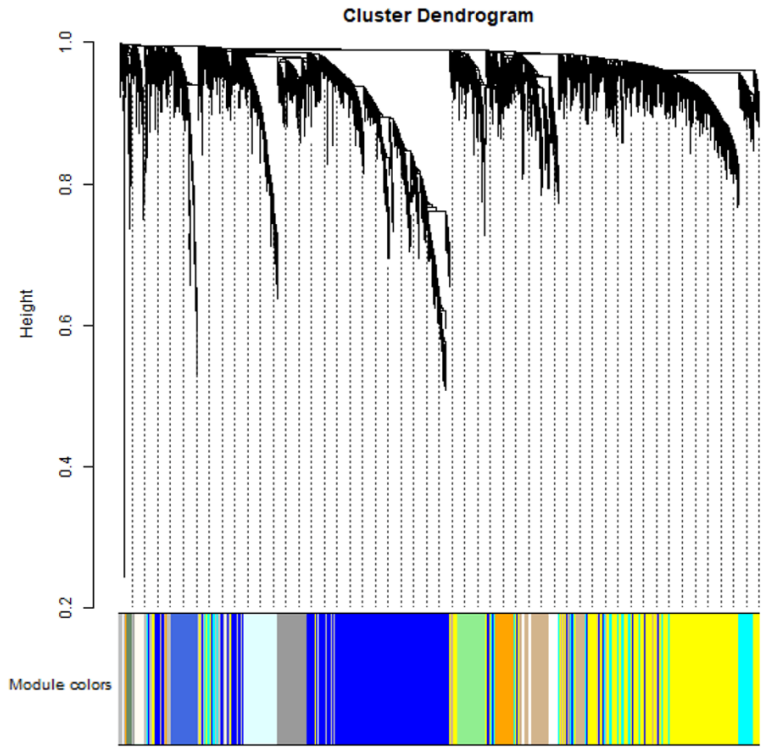

Figure 6. The cluster dendrogram and color display of co-expression network modules for aberrant genes from WGCNA.

tial ethnic difference. For genes located in the IL17 signaling pathway, S100A8, S100A9, and IL17F were downregulated in the TCGA dataset while S100A7, S100A7A, and IL1B were upregulated in the GSE142083 dataset.

When enriched in KEGG, 6 pathways were overlapped with GSE142083, including ECM-receptor interaction, protein digestion and absorption, amoebiasis, AGE-RAGE signaling pathway in diabetic complications, salivary secretion, and cAMP signaling pathway; and the first 4 pathways were upregulated. Besides, extracellular structure organization, extracellular matrix organization, and O-glycan processing were overlapped in GO with the first 2 being upregulated. The relatively small overlap of pathways also indicated the existence of racial disparities.

Furthermore, the IL-17 signaling pathway, the well-established pathway in Chinese samples

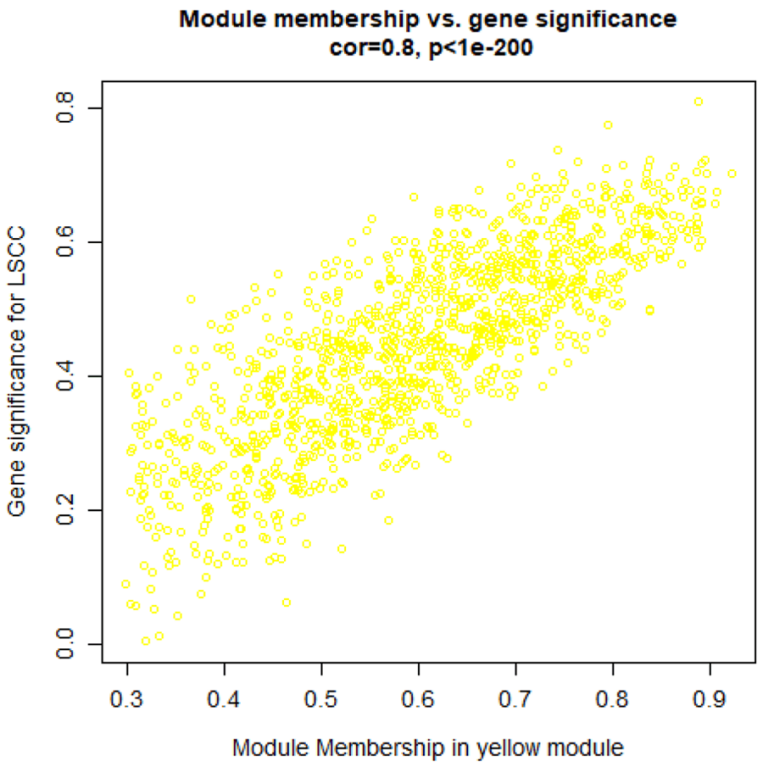

Figure 7. The scatterplot of correlation between module membership and gene significance in yellow module.

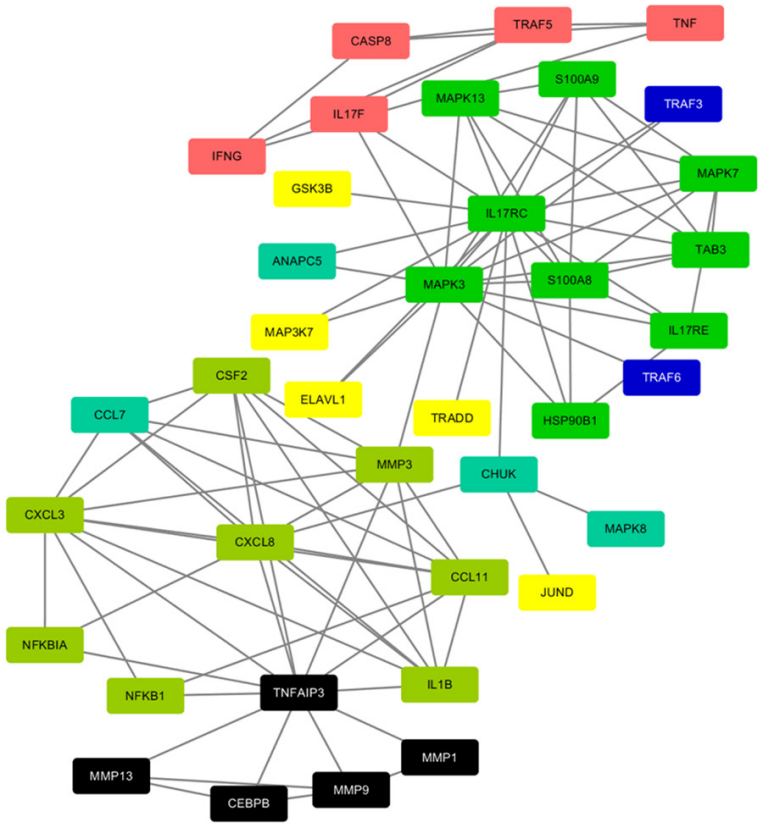

Figure 8. Hub genes located in IL-17 signaling pathway and different colors represent different modules.

from the abovementioned analyses, seemed not significantly enriched in the Caucasians; and epidermis development and epidermal cell differentiation pathways were downregulated, opposite to that in the Chinese (Figure S6). These may account for the ethnicity-specific effect. Additionally, it should be pointed out that MMP1 was still differentially expressed in the Caucasians while CXCL10 not. High IL17RC expression in LSCC tissue

IHC staining results

Finally, we applied IHC staining to further validate the protein expression level of IL17RC, a novel hub gene associated with LSCC. According to the results, 16 of 18 LSCC tissues were stained positively and only 1 normal tissue was stained positively, indicating a strong statistical significance (Fisher *p*-value <0.001). In summary, we validated the protein expression in 18 pairs of LSCC tissues and adjacent tissues by IHC staining and found significantly elevated LSCC expression in terms of density and intensity in LSCC tissues compared with adjacent tissues (**Figure 9A-D**).

Discussion

In this integrative analysis, we unveiled the key role of the IL-17 signaling pathway in LSCC and identified the hub genes closely associated with LSCC in this pathway. Furthermore, we evaluated the

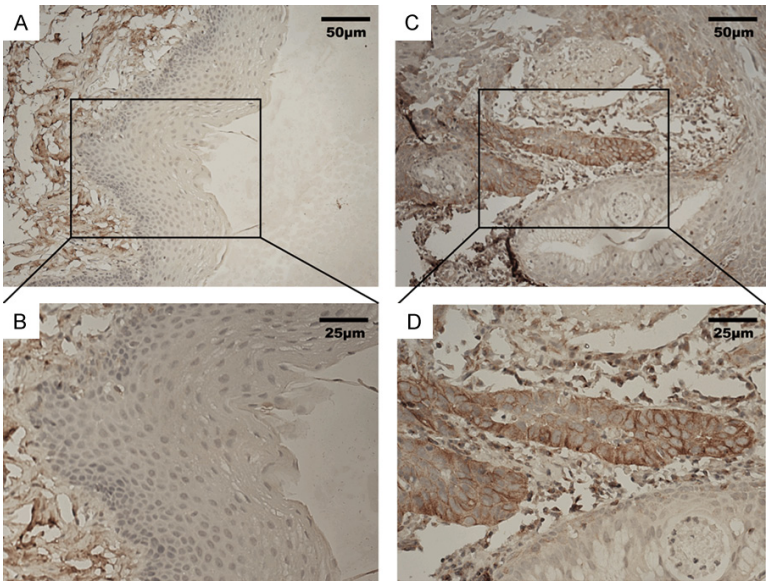

Figure 9. Representative immunohistochemical staining for IL17RC protein in 18 pairs of LSCC tissues and matched adjacent normal tissues. A, B. Negative expression of IL17RC in adjacent normal tissues (200×, 400×); C, D. Positive expression of IL17RC protein in LSCC specimens (200×, 400×).

ethnicity-specific effect of gene expression in LSCC. The hub gene L17RC was further verified to be upregulated in LSCC tissue using IHC staining.

As per the results, 1184 genes (553 upregulated and 595 downregulated) were differentially expressed in the discovery stage and two genes participating in the IL-17 signaling pathway (MMP1 and CXCL10) were further validated in GSE117005. Subsequently, only the IL-17 signaling pathway was validated as well. In WGCNA for GSE142083, the yellow module was the most significant one of the modules that were positively correlated with LSCC. When enriching genes in this module, many well-established carcinogenic pathways were brought out, such as the cell cycle, p53 signaling pathway, and nuclear division [19]. Furthermore, a PPI network was constructed for genes located in the IL-17 signaling pathway, and IL17RC, MAPK3, S100A8, MMP3, CXCL8, and TNFAIP3 were indicated to be hub genes. In comparison with results derived from Chinese, the IL-17 signaling pathway was not significantly enriched in TCGA datasets and the epidermis development pathway was in the opposite direction, suggesting a different expression profile of genes in Chinese and Caucasians.

In our study, the identified genes (MMP1 and MMP3) belong to the matrix metalloproteases family, participating in the breakdown of extracellular matrix and metastasis [20]. They were both upregulated in analyzed datasets. Meanwhile, MMP1 was confirmed to be over-expressed in LSCC by recent RNA-Seq studies [21]; and MMP3 was reported to be closely associated with LSCC and also related to tumor stage [22, 23]. These studies were all performed in Chinese and added integrity and accuracy to our study. S100A8, also known as calprotectin if combined with S100A9, was reported to be upregulated in HNSCC [24], together with CXCL8 [25]. CXC subfamily, including CXCL8 and CXCL10, has been reported to play an important role in regulating tumor growth and metastasis [26] while whether it affects LSCC is still unknown. In our study, CXCL8 and CXCL10 could involve in the initiation and progression of LSCC.

Although no direct association between IL17RC and LSCC or HNSCC was observed in previous publications, it was clear that IL17RC might affect the initiation and progression of other cancers [27]. Considering IL-17 is a pro-inflammatory factor, causing inflammation and promoting the initiation of carcinogenesis, it is reasonable that it plays a key role in the initiation and progression of LSCC. Since the IL-17 signaling pathway may be ethnicity-specific in LSCC, why this pathway was not enriched previously may be attributed to it that the main samples were Caucasians where this pathway was not significant. TNFAIP3 is a key gene suppressing inflammation, immunity, and carcinogenesis and affects many diseases [28]. Here, its logFC was more than 0 and *p*-value less than 0.01 in both GSE142083 and TCGA datasets, indicating a controversy to the function of TNFAIP3. We postulated that the expression of TNFAIP3 was in a negative feedback loop where the carcinogenesis induced the TNFAIP3 expression to suppress tumor initiation. As for

MAPK3, it involves in MAPK/ERK cascade pathway [29] and inhibition of it suppresses cancer-stromal interaction and metastasis [30]. However, its expression was slightly downregulated in these analyzed datasets, quite opposite to the knowns. Thus, it may be due to a negative feedback loop in LSCC where such cascade pathway is inactivated to combat the progression of cancer just the same as that of TNFAIP3.

The difference in mRNA expression profiles between Chinese and Caucasian samples is rather interesting. IL-17 signaling pathway has only been reported to have an ethnicity-specific effect in psoriasis and this pathway was also upregulated in Asian samples while downregulated in the Europeans [31]. We hypothesized that such different patterns might be caused by the genetic variability in different ethnicities. However, only two genome-wide association studies (GWAS) was performed on LSCC for the Chinese and Europeans respectively with both sample size less than 10,000 [32, 33]; and no trans-ethnic GWAS was available. Although genes located in the IL-17 signaling pathway were not significantly associated with LSCC in these two GWAS, the results indicated a substantial difference, suggesting the importance of genetic variability in LSCC and mRNA expression profile may be affected by it. Further trans-ethnic GWAS in large samples should be performed to unveil the underlying mechanism. As for the opposite regulation direction in the Chinese and Caucasians including epidermis development and epidermal cell differentiation, we also presume the genetic factors may contribute to it since no related studies have found it.

Our study unveiled the novel association between the IL-17 signaling pathway and LSCC; it also revealed the ethnic specificity of this pathway. The results suggested that immunotherapies targeting hub genes in this pathway might be effective in LSCC and racial disparities should be accounted for in clinical practice.

Conclusion

IL-17 signaling pathway plays a key role in LSCC, especially in Chinese; and hub genes located in it may be potential immunotherapeutic targets for LSCC.

Acknowledgements

A.Y contributed to the study design, supervised the data analysis process, and mainly revised the manuscript. L.Q and W.B were responsible for data acquisition, bioinformatic analysis, and data visualization; and L.Q drafted the original manuscript. WL and XD read and revised the original manuscript and gave substantial suggestions on statistics. All authors have approved the publication of this study. A.Y takes responsibility for the integrity of the data and the accuracy of the data analysis.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Aihui Yan, Department of Otorhinolaryngology, The First Hospital of China Medical University, 155 Nanjing Street, Heping District, Shenyang 110001, Liaoning Province, China. Tel: +86-13889323421; E-mail: yah012@sina.com

References

[1] Steuer CE, El-Deiry M, Parks JR, Higgins KA and Saba NF. An update on larynx cancer. CA Cancer J Clin 2017; 67: 31-50.

[2] Bonomi MR, Blakaj A and Blakaj D. Organ preservation for advanced larynx cancer: a review of chemotherapy and radiation combination strategies. Oral Oncol 2018; 86: 301-306.

[3] Solomon B, Young RJ and Rischin D. Head and neck squamous cell carcinoma: genomics and emerging biomarkers for immunomodulatory cancer treatments. Semin Cancer Biol 2018; 52: 228-240.

[4] Gao W, Zhang Y, Luo H, Niu M, Zheng X, Hu W, Cui J, Xue X, Bo Y, Dai F, Lu Y, Yang D, Guo Y, Guo H, Li H, Zhang Y, Yang T, Li L, Zhang L, Hou R, Wen S, An C, Ma T, Jin L, Xu W and Wu Y. Targeting SKA3 suppresses the proliferation and chemoresistance of laryngeal squamous cell carcinoma via impairing PLK1-AKT axis-mediated glycolysis. Cell Death Dis 2020; 11: 919.

[5] Ward EM, Flowers CR, Gansler T, Omer SB and Bednarczyk RA. The importance of immunization in cancer prevention, treatment, and survivorship. CA Cancer J Clin 2017; 67: 398-410.

[6] Cohen EEW, Bell RB, Bifulco CB, Burtness B, Gillison ML, Harrington KJ, Le QT, Lee NY, Leidner R, Lewis RL, Licitra L, Mehanna H, Mell LK, Raben A, Sikora AG, Uppaluri R, Whitworth F, Zandberg DP and Ferris RL. The society for

immunotherapy of cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC). J Immunother Cancer 2019; 7: 184.
[7] Li X, Bechara R, Zhao J, McGeachy MJ and Gaffen SL. IL-17 receptor-based signaling and implications for disease. Nat Immunol 2019; 20: 1594-1602.
[8] Coffelt SB, Kersten K, Doornebal CW, Weiden J, Vrijland K, Hau CS, Verstegen NJM, Ciampricotti M, Hawinkels LJAC, Jonkers J and de Visser KE. IL-17-producing γδ T cells and neutrophils conspire to promote breast cancer metastasis. Nature 2015; 522: 345-348.
[9] Wang X, Yang L, Huang F, Zhang Q, Liu S, Ma L and You Z. Inflammatory cytokines IL-17 and TNF-α up-regulate PD-L1 expression in human prostate and colon cancer cells. Immunol Lett 2017; 184: 7-14.
[10] Zhang Q, Liu S, Ge D, Cunningham DM, Huang F, Ma L, Burris TP and You Z. Targeting Th17-IL-17 pathway in prevention of micro-invasive prostate cancer in a mouse model. Prostate 2017; 77: 888-899.
[11] Gelbard A, Katsantonis NG, Mizuta M, Newcomb D, Rotsinger J, Rousseau B, Daniero JJ, Edell ES, Ekbom DC, Kasperbauer JL, Hillel AT, Yang L, Garrett CG, Netterville JL, Wootten CT, Francis DO, Stratton C, Jenkins K, McGregor TL, Gaddy JA, Blackwell TS and Drake WP. Idiopathic subglottic stenosis is associated with activation of the inflammatory IL-17A/IL-23 axis. Laryngoscope 2016; 126: E356-E361.
[12] Li FJ, Cai ZJ, Yang F, Zhang SD and Chen M. Th17 expression and IL-17 levels in laryngeal squamous cell carcinoma patients. Acta Otolaryngol 2016; 136: 484-490.
[13] Zhao R, Li FQ, Tian LL, Shang DS, Guo Y, Zhang JR and Liu M. Comprehensive analysis of the whole coding and non-coding RNA transcriptome expression profiles and construction of the circRNA-lncRNA co-regulated ceRNA network in laryngeal squamous cell carcinoma. Funct Integr Genomics 2019; 19: 109-121.
[14] Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, Brozell A, Schadt EE, Drake TA, Lusis AJ and Horvath S. Integrating genetic and network analysis to characterize genes related to mouse weight. PLoS Genet 2006; 2: e130.
[15] Kanehisa M, Furumichi M, Tanabe M, Sato Y and Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 2017; 45: D353-D361.
[16] The Gene Ontology Consortium. The gene ontology resource: 20 years and still GOing strong. Nucleic Acids Res 2019; 47: D330-D338.
[17] Robinson MD, McCarthy DJ and Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010; 26: 139-140.
[18] Yu G, Wang LG, Han Y and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.
[19] Lin AB, McNeely SC and Beckmann RP. Achieving precision death with cell-cycle inhibitors that target DNA replication and repair. Clin Cancer Res 2017; 23: 3232-3240.
[20] Cui N, Hu M and Khalil RA. Biochemical and biological attributes of matrix metalloproteinases. Prog Mol Biol Transl Sci 2017; 147: 1-73.
[21] Fang W and Shen J. Identification of MMP1 and MMP2 by RNA-seq analysis in laryngeal squamous cell carcinoma. Am J Otolaryngol 2020; 41: 102391.
[22] Hui L, Yang N, Yang H, Guo X and Jang X. Identification of biomarkers with a tumor stage-dependent expression and exploration of the mechanism involved in laryngeal squamous cell carcinoma. Oncol Rep 2015; 34: 2627-2635.
[23] Lapa RML, Barros-Filho MC, Marchi FA, Domingues MAC, de Carvalho GB, Drigo SA, Kowalski LP and Rogatto SR. Integrated miRNA and mRNA expression analysis uncovers drug targets in laryngeal squamous cell carcinoma patients. Oral Oncol 2019; 93: 76-84.
[24] Argyris PP, Slama ZM, Ross KF, Khammanivong A and Herzberg MC. Calprotectin and the initiation and progression of head and neck cancer. J Dent Res 2018; 97: 674-682.
[25] Michiels K, Schutyser E, Conings R, Lenaerts JP, Put W, Nuyts S, Delaere P, Jacobs R, Struyf S, Proost P and Van Damme J. Carcinoma cell-derived chemokines and their presence in oral fluid. Eur J Oral Sci 2009; 117: 362-368.
[26] Tokunaga R, Zhang W, Naseem M, Puccini A, Berger MD, Soni S, McSkane M, Baba H and Lenz HJ. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - a target for novel cancer therapy. Cancer Treat Rev 2018; 63: 40-47.
[27] Yan C, Huang WY, Boudreau J, Mayavannan A, Cheng Z and Wang J. IL-17R deletion predicts high-grade colorectal cancer and poor clinical outcomes. Int J Cancer 2019; 145: 548-558.
[28] Verstrepen L, Verhelst K, van Loo G, Carpentier I, Ley SC and Beyaert R. Expression, biological activities and mechanisms of action of A20 (TNFAIP3). Biochem Pharmacol 2010; 80: 2009-2020.
[29] Yang L, Zheng L, Chng WJ and Ding JL. Comprehensive analysis of ERK1/2 substrates for potential combination immunotherapies. Trends Pharmacol Sci 2019; 40: 897-910.

[30] Yan Z, Ohuchida K, Fei S, Zheng B, Guan W, Feng H, Kibe S, Ando Y, Koikawa K, Abe T, Iwamoto C, Shindo K, Moriyama T, Nakata K, Miyasaka Y, Ohtsuka T, Mizumoto K, Hashizume M and Nakamura M. Inhibition of ERK1/2 in cancer-associated pancreatic stellate cells suppresses cancer-stromal interaction and metastasis. J Exp Clin Cancer Res 2019; 38: 221.

[31] Kim J, Oh CH, Jeon J, Baek Y, Ahn J, Kim DJ, Lee HS, Correa da Rosa J, Suárez-Fariñas M, Lowes MA and Krueger JG. Molecular phenotyping small (Asian) versus large (Western) plaque psoriasis shows common activation of IL-17 pathway genes but different regulatory gene sets. J Invest Dermatol 2016; 136: 161-172.

[32] Wei Q, Yu D, Liu M, Wang M, Zhao M, Liu M, Jia W, Ma H, Fang J, Xu W, Chen K, Xu Z, Wang J, Tian L, Yuan H, Chang J, Hu Z, Wei L, Huang Y, Han Y, Liu J, Han D, Shen H, Yang S, Zheng H, Ji Q, Li D, Tan W, Wu C and Lin D. Genome-wide association study identifies three susceptibility loci for laryngeal squamous cell carcinoma in the Chinese population. Nat Genet 2014; 46: 1110-1114.

[33] Shete S, Liu H, Wang J, Yu R, Sturgis EM, Li G, Dahlstrom KR, Liu Z, Amos CI and Wei Q. A genome-wide association study identifies two novel susceptible regions for squamous cell carcinoma of the head and neck. Cancer Res 2020; 80: 2451-2460.

Table S1. Top 50 differentially expressed genes in the GSE142083 dataset

symbol	logFC	logCPM	LR	*P*Value	FDR
MMP1	5.573235	8.524383	309.6588	2.59E-69	3.93E-66
SPRR2G	5.281806	8.663347	142.0687	9.39E-33	9.43E-31
PAGE2	4.900713	3.794749	315.153	1.65E-70	3.05E-67
PRR9	4.807902	5.573755	169.434	9.84E-39	2.00E-36
S100A7A	4.780252	7.810139	89.2451	3.49E-21	1.07E-19
LCE1F	4.687459	4.894621	111.2125	5.32E-26	2.71E-24
DEFA5	4.648037	3.868147	325.7899	7.94E-73	1.89E-69
MAGEA3	4.599776	5.236519	104.8149	1.34E-24	5.99E-23
HOXD11	4.504224	5.546649	147.4871	6.14E-34	6.96E-32
MAGEC2	4.436644	3.750476	277.0409	3.31E-62	3.25E-59
SOX11	4.406962	4.558045	418.3095	5.69E-93	4.74E-89
MAGEA6	4.348905	5.040277	77.46377	1.35E-18	3.11E-17
MAGEA4	4.284282	3.826733	287.1872	2.04E-64	2.61E-61
MAGEC1	4.276632	3.463943	190.1792	2.91E-43	1.05E-40
FTHL17	4.178194	3.370975	165.283	7.93E-38	1.41E-35
LCE2A	4.06795	4.128031	85.75332	2.04E-20	5.64E-19
LCE3D	4.042338	8.391843	63.53308	1.58E-15	2.62E-14
MAGEA1	4.014648	3.327204	149.182	2.62E-34	3.07E-32
MAGEA10	3.991598	3.613637	210.1878	1.25E-47	6.94E-45
LCE1A	3.978382	4.405897	61.95839	3.51E-15	5.59E-14
NEU2	3.949829	3.547017	191.6608	1.38E-43	5.11E-41
MMP11	3.931822	5.053038	449.9864	7.26E-100	1.21E-95
LCE3A	3.928341	6.559446	60.49925	7.36E-15	1.12E-13
GAGE1	3.920222	3.248169	132.2948	1.29E-30	1.05E-28
S100A7	3.858901	10.23632	55.11131	1.14E-13	1.53E-12
C1orf68	3.858095	4.18274	91.35416	1.20E-21	3.92E-20
DLX2	3.85262	4.248706	276.0518	5.45E-62	5.04E-59
KRT75	3.766844	4.993354	90.26211	2.09E-21	6.63E-20
LCE2C	3.765898	4.004631	64.76211	8.45E-16	1.44E-14
CASP14	3.747534	4.515179	94.9634	1.94E-22	6.88E-21
KHDC1L	3.705634	4.805561	138.6392	5.28E-32	4.97E-30
HOXD13	3.671757	4.20341	112.4158	2.90E-26	1.52E-24
OLFM4	3.652963	3.548935	177.3296	1.86E-40	4.69E-38
PYDC2	3.64813	4.742767	78.5426	7.83E-19	1.87E-17
NLRP10	3.594602	3.507927	165.3287	7.75E-38	1.39E-35
MMP3	3.576477	5.929729	133.0645	8.75E-31	7.22E-29
KRT16	3.548703	11.11783	63.27396	1.80E-15	2.97E-14
PRAME	3.547824	4.072016	127.9525	1.15E-29	8.37E-28
CCL11	3.544823	4.207391	112.3396	3.01E-26	1.57E-24
DSC1	3.531399	4.895296	77.24593	1.51E-18	3.46E-17
CA9	3.519589	3.90249	229.7401	6.79E-52	5.15E-49
LCE6A	3.514122	3.62457	74.62456	5.69E-18	1.23E-16
CSF2	3.508788	3.785118	223.4607	1.59E-50	1.10E-47
KRTDAP	3.504813	8.427482	45.40744	1.60E-11	1.67E-10
CALB1	3.469184	4.369388	110.731	6.78E-26	3.39E-24
SPRR2F	3.461568	7.670836	47.48225	5.55E-12	6.13E-11
SPINK6	3.456887	3.444787	144.2452	3.14E-33	3.31E-31
HOXC10	3.443711	3.697596	110.4705	7.73E-26	3.84E-24
MMP13	3.420874	5.997841	126.3693	2.55E-29	1.82E-27

SPRR2E	3.411051	7.534786	56.51957	5.56E-14	7.72E-13
C6orf58	-6.02891	8.91495	119.7367	7.22E-28	4.36E-26
PIP	-5.98293	9.328714	135.0599	3.20E-31	2.78E-29
PRB3	-5.94591	12.13987	112.0102	3.55E-26	1.84E-24
PRH1	-5.85551	14.39301	99.59714	1.87E-23	7.32E-22
PRH2	-5.82857	12.81	107.3808	3.67E-25	1.71E-23
STATH	-5.82053	10.06336	105.8089	8.12E-25	3.69E-23
BPIFB2	-5.64145	7.483564	141.9748	9.85E-33	9.74E-31
CRISP3	-5.63874	7.680236	209.9783	1.39E-47	7.47E-45
SCGB3A1	-5.46014	12.34463	105.8622	7.90E-25	3.60E-23
LPO	-5.43159	6.080222	162.0619	4.01E-37	6.37E-35
PIGR	-5.31071	11.15338	121.3483	3.21E-28	2.02E-26
GP2	-5.30182	6.925129	174.5702	7.43E-40	1.65E-37
ZG16B	-5.29934	9.44134	122.2223	2.06E-28	1.32E-26
LRRC26	-5.1961	7.986973	149.4883	2.24E-34	2.69E-32
BPIFB1	-5.19185	10.22879	95.64261	1.38E-22	4.93E-21
MUC5B	-5.1308	10.24395	96.97732	7.01E-23	2.55E-21
AQP5	-5.08755	8.652578	198.1896	5.19E-45	2.11E-42
KRT4	-4.84445	12.62044	52.74958	3.79E-13	4.82E-12
ADH1B	-4.59239	4.919417	342.7789	1.58E-76	4.40E-73
LTF	-4.55823	9.439858	61.49654	4.44E-15	6.95E-14
DMBT1	-4.47923	7.546079	83.53246	6.27E-20	1.66E-18
MUC5AC	-4.44475	7.142114	99.9398	1.57E-23	6.22E-22
SCGB3A2	-4.43556	8.681597	60.86825	6.10E-15	9.36E-14
MUC19	-4.38336	4.036936	286.5059	2.87E-64	3.42E-61
LYZ	-4.37115	14.31177	89.3385	3.33E-21	1.03E-19
AZGP1	-4.31986	8.856974	65.34175	6.30E-16	1.10E-14
MPO	-4.28339	4.438956	150.7325	1.20E-34	1.54E-32
PRB1	-4.26966	4.630803	128.3551	9.39E-30	6.89E-28
PRB4	-4.25023	9.153439	88.87426	4.21E-21	1.28E-19
TFF3	-4.22368	7.662691	104.6409	1.46E-24	6.51E-23
MUC21	-4.16077	9.226288	41.52412	1.16E-10	1.10E-09
ADIPOQ	-4.09228	3.799308	145.4085	1.75E-33	1.87E-31
FOLR1	-3.91537	5.88509	120.4243	5.11E-28	3.14E-26
SLC13A2	-3.88441	3.860638	253.1087	5.45E-57	4.78E-54
WFDC2	-3.87366	7.781231	114.5718	9.77E-27	5.35E-25
FAM3D	-3.84924	5.892937	414.1836	4.50E-92	2.50E-88
BPIFA1	-3.83696	7.169477	57.34322	3.66E-14	5.19E-13
MYOC	-3.82475	3.894488	89.69716	2.78E-21	8.68E-20
TFF1	-3.81968	4.328133	118.7431	1.19E-27	7.10E-26
GSTA2	-3.73577	3.488136	168.9509	1.25E-38	2.46E-36
MAL	-3.6955	6.73779	38.90146	4.46E-10	3.89E-09
PAX1	-3.68366	4.201087	114.8091	8.66E-27	4.80E-25
CRNN	-3.64531	10.67294	21.1693	4.20E-06	2.07E-05
ODAM	-3.63733	5.5773	48.6548	3.05E-12	3.47E-11
MS4A8	-3.62862	4.184844	140.7513	1.82E-32	1.79E-30
CST5	-3.58736	3.222383	114.1372	1.22E-26	6.62E-25
TMPRSS11B	-3.57446	7.243999	29.57775	5.37E-08	3.60E-07
CLCA4	-3.52345	7.138373	88.708	4.58E-21	1.39E-19
KRT13	-3.50645	13.78897	35.54337	2.49E-09	1.99E-08
MUC7	-3.48056	4.442312	66.32582	3.82E-16	6.81E-15

Note: “symbol” is gene symbol; “logFC” is log fold change; “logCPM” is log counts per million; “LR” is likelihood ratio; “*P*Value” is the original *p*-value for logFC; “FDR” is adjusted *p*-value for logFC.

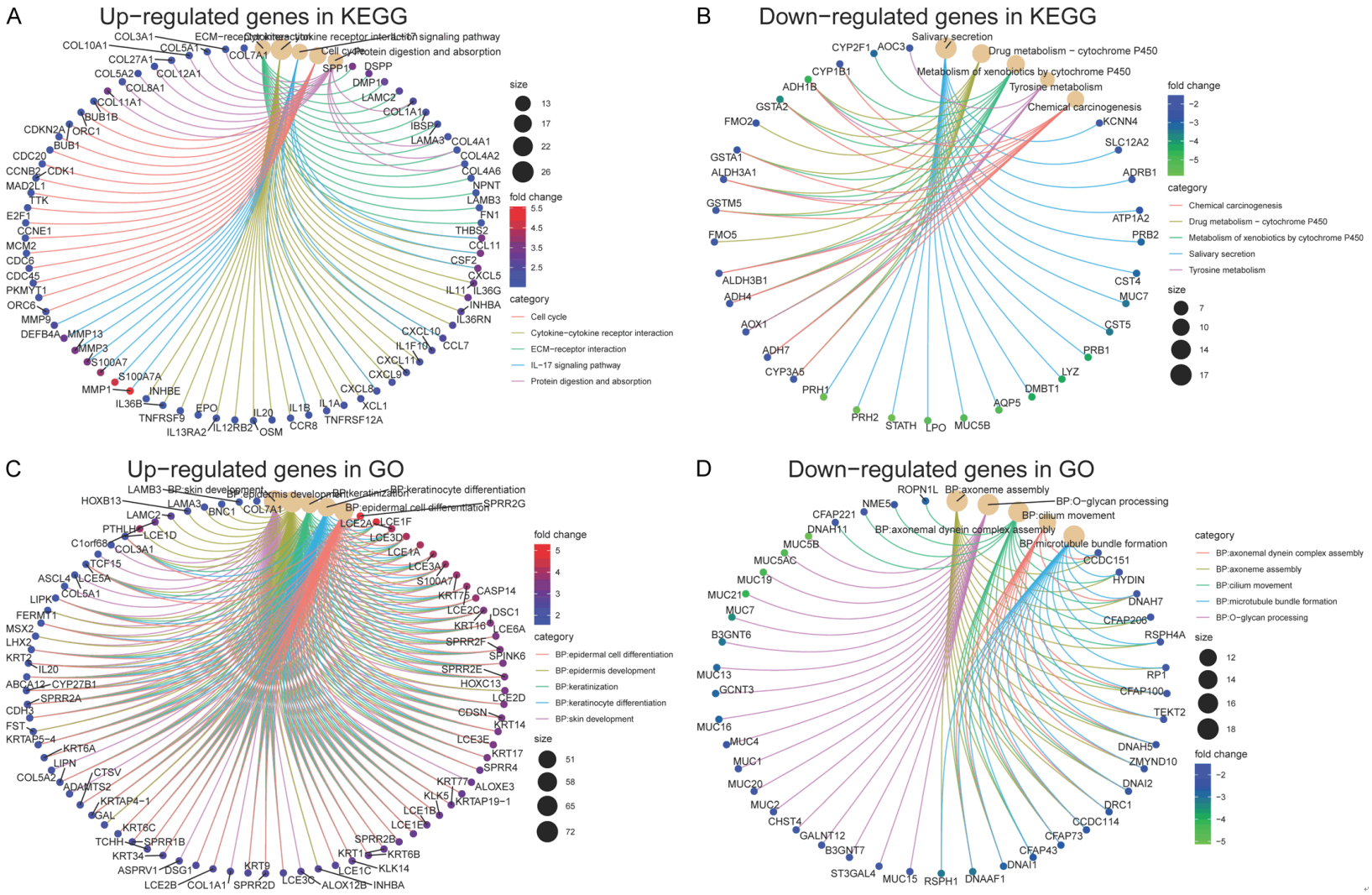

Figure S1. Network of connection between gene and pathway in GSE142083.

Table S2. Top 50 enriched pathways in the GSE142083 dataset

ID	Description	GeneRatio	BgRatio	*p*value	qvalue	Count	type
hsa04970	Salivary secretion	17/224	93/8081	4.90E-10	1.03E-07	17	DOWN
hsa04512	ECM-receptor interaction	14/225	88/8081	1.16E-07	1.57E-05	14	UP
hsa04060	Cytokine-cytokine receptor interaction	26/225	295/8081	1.59E-07	1.57E-05	26	UP
hsa04657	IL-17 signaling pathway	14/225	94/8081	2.72E-07	1.57E-05	14	UP
hsa04110	Cell cycle	16/225	124/8081	2.89E-07	1.57E-05	16	UP
hsa00982	Drug metabolism - cytochrome P450	12/224	72/8081	5.42E-07	5.68E-05	12	DOWN
hsa04657	IL-17 signaling pathway	19/449	94/8081	7.20E-07	0.000196	19	ALL
hsa04060	Cytokine-cytokine receptor interaction	37/449	295/8081	2.31E-06	0.000249	37	ALL
hsa04970	Salivary secretion	18/449	93/8081	2.74E-06	0.000249	18	ALL
hsa04974	Protein digestion and absorption	13/225	103/8081	4.98E-06	0.000216	13	UP
hsa00980	Metabolism of xenobiotics by cytochrome P450	11/224	78/8081	8.82E-06	0.000616	11	DOWN
hsa05146	Amoebiasis	12/225	102/8081	2.39E-05	0.000865	12	UP
hsa04974	Protein digestion and absorption	17/449	103/8081	4.60E-05	0.003138	17	ALL
hsa00350	Tyrosine metabolism	7/224	36/8081	4.81E-05	0.002521	7	DOWN
hsa04512	ECM-receptor interaction	15/449	88/8081	8.77E-05	0.00478	15	ALL
hsa05204	Chemical carcinogenesis	10/224	83/8081	9.07E-05	0.003798	10	DOWN
hsa05222	Small cell lung cancer	10/225	92/8081	0.000224	0.006945	10	UP
hsa05150	Staphylococcus aureus infection	15/449	96/8081	0.00024	0.010908	15	ALL
hsa04024	cAMP signaling pathway	16/224	216/8081	0.000316	0.011027	16	DOWN
hsa04061	Viral protein interaction with cytokine and cytokine receptor	10/225	100/8081	0.000443	0.010682	10	UP
hsa04933	AGE-RAGE signaling pathway in diabetic complications	10/225	100/8081	0.000443	0.010682	10	UP
hsa00982	Drug metabolism - cytochrome P450	12/449	72/8081	0.000545	0.021238	12	ALL
hsa03320	PPAR signaling pathway	8/224	76/8081	0.001145	0.034271	8	DOWN
hsa04061	Viral protein interaction with cytokine and cytokine receptor	14/449	100/8081	0.001185	0.039294	14	ALL
hsa04110	Cell cycle	16/449	124/8081	0.00133	0.039294	16	ALL
hsa05150	Staphylococcus aureus infection	9/225	96/8081	0.001357	0.029421	9	UP
hsa05146	Amoebiasis	14/449	102/8081	0.001441	0.039294	14	ALL
hsa00010	Glycolysis/Gluconeogenesis	7/224	67/8081	0.002424	0.063458	7	DOWN
hsa04261	Adrenergic signaling in cardiomyocytes	11/224	150/8081	0.002945	0.068556	11	DOWN
hsa04926	Relaxin signaling pathway	10/225	129/8081	0.003127	0.061645	10	UP
hsa00350	Tyrosine metabolism	7/449	36/8081	0.003194	0.07915	7	ALL
hsa00980	Metabolism of xenobiotics by cytochrome P450	11/449	78/8081	0.003664	0.083248	11	ALL
hsa04115	p53 signaling pathway	7/225	73/8081	0.004046	0.065882	7	UP
hsa00830	Retinol metabolism	10/449	68/8081	0.004047	0.084877	10	ALL

hsa04218	Cellular senescence	11/225	156/8081	0.004119	0.065882	11	UP
hsa05323	Rheumatoid arthritis	8/225	93/8081	0.004254	0.065882	8	UP
hsa05414	Dilated cardiomyopathy	8/224	96/8081	0.005024	0.091428	8	DOWN
hsa04216	Ferroptosis	5/224	41/8081	0.005201	0.091428	5	DOWN
hsa05219	Bladder cancer	5/225	41/8081	0.005299	0.076603	5	UP
hsa04978	Mineral absorption	6/224	59/8081	0.005612	0.091428	6	DOWN
hsa00533	Glycosaminoglycan biosynthesis - keratan sulfate	4/449	14/8081	0.006025	0.117336	4	ALL
hsa00533	Glycosaminoglycan biosynthesis - keratan sulfate	3/224	14/8081	0.006101	0.091428	3	DOWN
hsa00601	Glycosphingolipid biosynthesis - lacto and neolacto series	4/224	27/8081	0.006111	0.091428	4	DOWN
hsa05202	Transcriptional misregulation in cancer	12/225	192/8081	0.007272	0.09855	12	UP
hsa03320	PPAR signaling pathway	10/449	76/8081	0.008957	0.162792	10	ALL
hsa04114	Oocyte meiosis	9/225	129/8081	0.009755	0.124429	9	UP
hsa00360	Phenylalanine metabolism	3/224	17/8081	0.010721	0.138526	3	DOWN
hsa05165	Human papillomavirus infection	17/225	331/8081	0.010828	0.130437	17	UP
hsa00830	Retinol metabolism	6/224	68/8081	0.011105	0.138526	6	DOWN
hsa00512	Mucin type O-glycan biosynthesis	4/224	32/8081	0.011242	0.138526	4	DOWN
GO:0043588	BP: skin development	72/499	412/17913	6.61E-37	2.29E-33	72	UP
GO:0008544	BP: epidermis development	70/499	452/17913	2.04E-32	3.53E-29	70	UP
GO:0031424	BP: keratinization	51/499	224/17913	4.84E-32	5.58E-29	51	UP
GO:0030216	BP: keratinocyte differentiation	56/499	298/17913	1.81E-30	1.57E-27	56	UP
GO:0043588	BP: skin development	87/977	412/17913	1.55E-28	6.76E-25	87	ALL
GO:0009913	BP: epidermal cell differentiation	56/499	346/17913	5.28E-27	3.65E-24	56	UP
GO:0031424	BP: keratinization	61/977	224/17913	1.64E-26	3.59E-23	61	ALL
GO:0008544	BP: epidermis development	88/977	452/17913	3.52E-26	5.13E-23	88	ALL
GO:0030216	BP: keratinocyte differentiation	69/977	298/17913	2.82E-25	3.09E-22	69	ALL
GO:0070268	BP: cornification	31/499	112/17913	1.27E-22	7.33E-20	31	UP
GO:0009913	BP: epidermal cell differentiation	71/977	346/17913	1.13E-22	8.97E-20	71	ALL
GO:0070268	BP: cornification	40/977	112/17913	1.23E-22	8.97E-20	40	ALL
GO:0031012	BP: extracellular matrix	77/1010	468/18678	1.23E-18	4.86E-16	77	ALL
GO:0030198	BP: extracellular matrix organization	44/499	334/17913	7.77E-18	3.84E-15	44	UP
GO:0005930	CC: axoneme	25/519	100/18678	2.47E-17	5.06E-15	25	DOWN
GO:0097014	CC: ciliary plasm	25/519	101/18678	3.19E-17	5.06E-15	25	DOWN
GO:0043062	BP: extracellular structure organization	47/499	387/17913	1.52E-17	6.55E-15	47	UP
GO:0062023	CC: collagen-containing extracellular matrix	67/1010	399/18678	8.06E-17	1.59E-14	67	ALL
GO:0044420	CC: extracellular matrix component	18/491	47/18678	5.97E-17	1.93E-14	18	UP
GO:0005201	CC: extracellular matrix structural constituent	41/960	158/16969	7.52E-17	4.69E-14	41	ALL

GO:0044447	CC: axoneme part	15/519	31/18678	7.43E-16	7.85E-14	15	DOWN
GO:0098644	CC: complex of collagen trimers	12/491	18/18678	1.53E-15	2.49E-13	12	UP
GO:0030198	BP: extracellular matrix organization	59/977	334/17913	7.85E-16	4.90E-13	59	ALL
GO:0043062	BP: extracellular structure organization	64/977	387/17913	1.26E-15	6.86E-13	64	ALL
GO:0031012	CC: extracellular matrix	46/491	468/18678	1.30E-14	1.40E-12	46	UP
GO:0035082	BP: axoneme assembly	18/478	52/17913	6.34E-16	2.19E-12	18	DOWN
GO:0044420	CC: extracellular matrix component	20/1010	47/18678	9.21E-14	1.21E-11	20	ALL
GO:0016266	BP: O-glycan processing	18/478	60/17913	1.13E-14	1.42E-11	18	DOWN
GO:0003341	BP: cilium movement	17/478	52/17913	1.24E-14	1.42E-11	17	DOWN
GO:0000280	BP: nuclear division	40/499	357/17913	6.51E-14	2.50E-11	40	UP
GO:0140014	BP: mitotic nuclear division	32/499	237/17913	1.37E-13	4.76E-11	32	UP
GO:0044447	CC: axoneme part	16/1010	31/18678	6.64E-13	6.55E-11	16	ALL
GO:0007059	BP: chromosome segregation	34/499	275/17913	3.14E-13	9.87E-11	34	UP
GO:0062023	CC: collagen-containing extracellular matrix	39/491	399/18678	1.75E-12	1.42E-10	39	UP
GO:0005201	MF: extracellular matrix structural constituent	26/477	158/16969	3.17E-13	1.53E-10	26	UP
GO:0005930	BP: axoneme	27/1010	100/18678	1.89E-12	1.49E-10	27	ALL
GO:0097014	CC: ciliary plasm	27/1010	101/18678	2.45E-12	1.61E-10	27	ALL
GO:0001533	CC: cornified envelope	14/491	44/18678	3.50E-12	2.27E-10	14	UP
GO:0007389	BP: pattern specification process	61/977	408/17913	6.13E-13	2.98E-10	61	ALL
GO:0009952	BP: anterior/posterior pattern specification	27/499	182/17913	1.19E-12	3.44E-10	27	UP
GO:0070286	BP: axonemal dynein complex assembly	12/478	25/17913	4.32E-13	3.73E-10	12	DOWN
GO:0048285	BP: organelle fission	40/499	395/17913	1.76E-12	4.60E-10	40	UP
GO:0000070	BP: mitotic sister chromatid segregation	23/499	132/17913	1.86E-12	4.60E-10	23	UP
GO:0098644	CC: complex of collagen trimers	12/1010	18/18678	8.05E-12	4.54E-10	12	ALL
GO:0001501	BP: skeletal system development	44/499	479/17913	3.73E-12	8.61E-10	44	UP
GO:0001578	BP: microtubule bundle formation	18/478	78/17913	1.66E-12	1.15E-09	18	DOWN
GO:0000819	BP: sister chromatid segregation	24/499	157/17913	1.13E-11	2.44E-09	24	UP
GO:0032838	CC: plasma membrane bounded cell projection cytoplasm	25/519	181/18678	3.94E-11	3.12E-09	25	DOWN
GO:0003002	BP: regionalization	34/499	318/17913	1.95E-11	3.97E-09	34	UP
GO:0005581	CC: collagen trimer	17/491	86/18678	7.77E-11	4.20E-09	17	UP

Note: “ID” is pathway ID in KEGG or GO; “Description” is a description of the pathway where CC is cellular component, BP is biological process and MF is molecular function; “GeneRatio” is the proportion of genes located in the pathway; “BgRatio” is background ratio; “*p*value” is *p*-value of the enriched pathway; “qvalue” is q-value of enriched pathway: “geneID” is gene symbol; “Count” is the number of genes. “type” is genes in up/down/all.

Table S3. Genes differentially expressed in both GSE142083 and GSE117005 datasets

symbol	logFC	logCPM	LR	*P*Value	FDR	change	dataset
ALOXE3	0.589792	7.849327	13.08917	0.000297	0.040628	UP	GSE117005
ALOXE3	3.114066	3.547042	146.0019	1.30E-33	1.40E-31	UP	GSE142083
BARX2	-0.75772	7.724124	19.53084	9.90E-06	0.004567	DOWN	GSE117005
BARX2	-1.83547	3.868505	99.2961	2.17E-23	8.43E-22	DOWN	GSE142083
CLCA4	-0.58902	8.5289	21.36704	3.79E-06	0.002983	DOWN	GSE117005
CLCA4	-3.52345	7.138373	88.708	4.58E-21	1.39E-19	DOWN	GSE142083
CLEC3B	-0.60204	8.067765	16.02003	6.27E-05	0.015432	DOWN	GSE117005
CLEC3B	-2.39952	4.77893	146.4915	1.01E-33	1.10E-31	DOWN	GSE142083
CRNN	-0.55101	8.65585	20.51546	5.92E-06	0.003102	DOWN	GSE117005
CRNN	-3.64531	10.67294	21.1693	4.20E-06	2.07E-05	DOWN	GSE142083
CXCL10	0.603537	8.189387	17.51635	2.85E-05	0.008961	UP	GSE117005
CXCL10	2.471419	7.873185	44.23899	2.91E-11	2.95E-10	UP	GSE142083
FAM3D	-0.55605	8.18052	14.86895	0.000115	0.021754	DOWN	GSE117005
FAM3D	-3.84924	5.892937	414.1836	4.50E-92	2.50E-88	DOWN	GSE142083
GRIN2D	0.674585	7.572616	13.88964	0.000194	0.029891	UP	GSE117005
GRIN2D	2.166957	3.189331	51.55366	6.97E-13	8.64E-12	UP	GSE142083
HBB	-0.5098	8.482703	15.5639	7.98E-05	0.017109	DOWN	GSE117005
HBB	-2.17505	8.802518	99.13761	2.36E-23	9.09E-22	DOWN	GSE142083
HOXA13	0.666887	7.467829	12.5686	0.000392	0.046279	UP	GSE117005
HOXA13	2.41447	3.28777	71.54446	2.71E-17	5.43E-16	UP	GSE142083
HOXC10	0.789619	7.525199	18.19132	2.00E-05	0.007252	UP	GSE117005
HOXC10	3.443711	3.697596	110.4705	7.73E-26	3.84E-24	UP	GSE142083
HOXD13	0.697854	7.601095	15.14891	9.94E-05	0.020385	UP	GSE117005
HOXD13	3.671757	4.20341	112.4158	2.90E-26	1.52E-24	UP	GSE142083
LRRC15	0.622372	7.695807	12.99757	0.000312	0.040628	UP	GSE117005
LRRC15	3.061725	4.566277	186.0287	2.34E-42	7.50E-40	UP	GSE142083
LRRN4CL	-0.74641	7.588565	17.1702	3.42E-05	0.009487	DOWN	GSE117005
LRRN4CL	-1.95878	4.652414	149.6662	2.05E-34	2.49E-32	DOWN	GSE142083
MAGEA1	0.863155	7.698843	24.52533	7.33E-07	0.001085	UP	GSE117005
MAGEA1	4.014648	3.327204	149.182	2.62E-34	3.07E-32	UP	GSE142083
MMP1	0.689931	8.051566	20.58085	5.72E-06	0.003102	UP	GSE117005
MMP1	5.573235	8.524383	309.6588	2.59E-69	3.93E-66	UP	GSE142083
MUC21	-1.07556	7.890258	24.1629	8.85E-07	0.001085	DOWN	GSE117005
MUC21	-4.16077	9.226288	41.52412	1.16E-10	1.10E-09	DOWN	GSE142083
PLN	-0.69117	7.510414	13.96003	0.000187	0.029891	DOWN	GSE117005
PLN	-1.54606	3.849454	62.30109	2.95E-15	4.76E-14	DOWN	GSE142083
PPP1R3C	-0.51397	8.214847	13.0547	0.000303	0.040628	DOWN	GSE117005
PPP1R3C	-2.07637	6.486529	99.16046	2.33E-23	9.00E-22	DOWN	GSE142083
RERGL	-1.02767	7.419901	27.93317	1.26E-07	0.000593	DOWN	GSE117005
RERGL	-2.42878	3.042429	48.31645	3.63E-12	4.09E-11	DOWN	GSE142083
SH3BGRL2	-0.61909	8.037027	14.51282	0.000139	0.025266	DOWN	GSE117005
SH3BGRL2	-2.73383	4.963331	282.1765	2.52E-63	2.62E-60	DOWN	GSE142083
SLURP1	-0.55778	8.429138	13.50139	0.000238	0.035155	DOWN	GSE117005
SLURP1	-1.82625	7.704144	7.588022	0.005876	0.015159	DOWN	GSE142083
SPRR2G	0.703401	7.795864	17.74328	2.53E-05	0.008521	UP	GSE117005
SPRR2G	5.281806	8.663347	142.0687	9.39E-33	9.43E-31	UP	GSE142083
SPRR4	0.668552	7.594391	13.86558	0.000196	0.029891	UP	GSE117005

SPRR4	3.115927	5.056379	75.42205	3.80E-18	8.39E-17	UP	GSE142083
TMPRSS11B	-0.7792	8.23449	20.94453	4.73E-06	0.003102	DOWN	GSE117005
TMPRSS11B	-3.57446	7.243999	29.57775	5.37E-08	3.60E-07	DOWN	GSE142083
XCL1	0.691811	7.516263	13.9936	0.000183	0.029891	UP	GSE117005
XCL1	2.02593	3.719746	97.27051	6.05E-23	2.22E-21	UP	GSE142083

Note: "symbol" is gene symbol; "logFC" is log fold change; "logCPM" is log counts per million; "LR" is likelihood ratio; "*P*Value" is orginal *p*-value for logFC; "FDR" is adjusted *p*-value for logFC; "change" is whether it is upregulated or downregulated; "dataset" is the analyzed dataset.

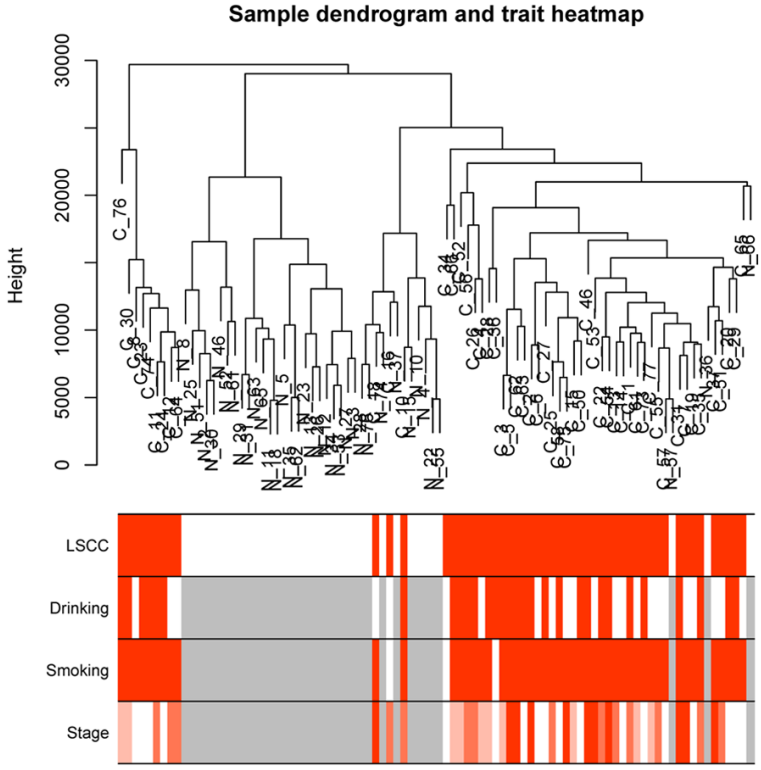

Figure S2. Cluster analysis for samples in GSE142083.

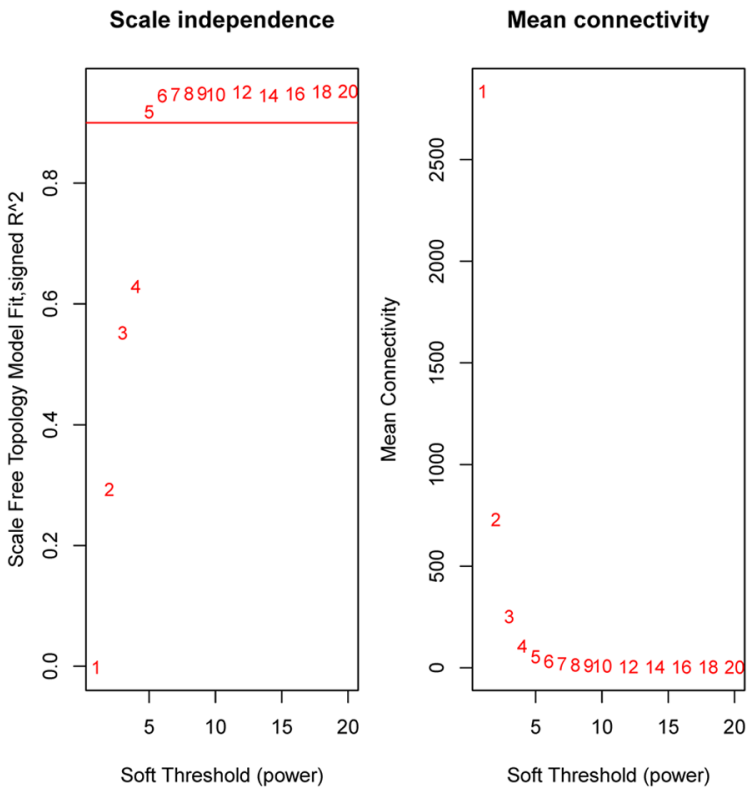

Figure S3. Analysis of network topology for soft-thresholding powers in GSE142083. The left panel shows the scale-free fit index (y-axis) as a function of the soft-thresholding power (x-axis). The right panel displays the mean connectivity (degree, y-axis) as a function of the soft-thresholding power (x-axis).

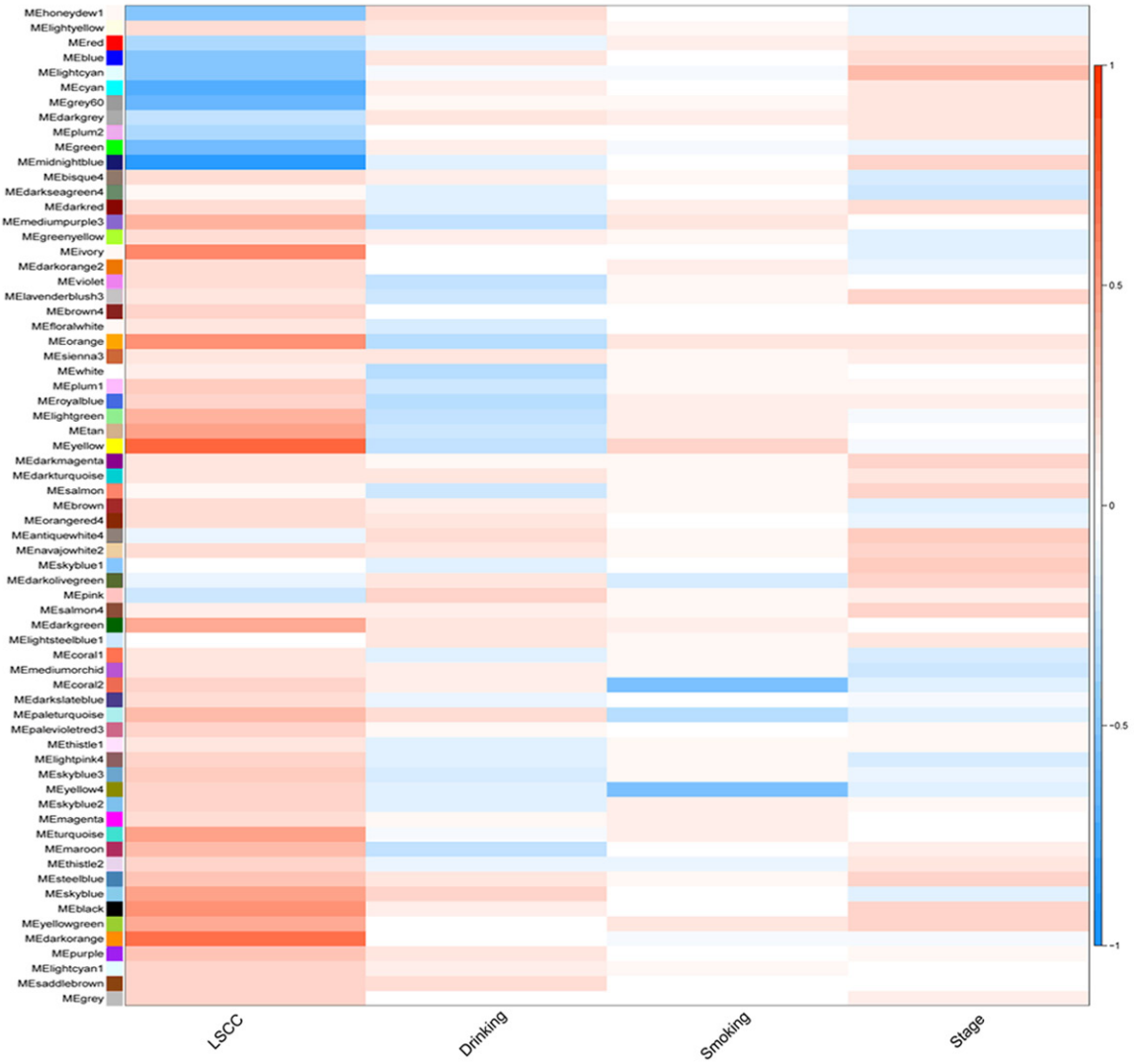

Figure S4. Heatmap of the relationship between modules (y-axis) and traits (x-axis) of GSE142083. Different colors represent different correlation coefficients.

Table S4. Significant genes in both the yellow module and the GSE1142083 dataset

geneSymbol	GS.LSCC	p.GS.LSCC
GINS1	0.723842543	7.61E-16
MYBL2	0.719703669	1.32E-15
CDC6	0.716735687	1.96E-15
HOXB7	0.715200627	2.39E-15
CDCA5	0.712701967	3.31E-15
SHCBP1	0.702854525	1.15E-14
KIF2C	0.702702823	1.17E-14
NCAPH	0.702618997	1.18E-14
KIF4A	0.70246615	1.20E-14
CENPA	0.698557396	1.94E-14
DTL	0.690589896	5.03E-14

ORC6	0.688907235	6.12E-14
AURKA	0.688651571	6.31E-14
KIF18B	0.686823036	7.80E-14
TEDC2	0.680517562	1.60E-13
KIF14	0.677814042	2.17E-13
BIRC5	0.674520353	3.13E-13
WDHD1	0.674517887	3.13E-13
AUNIP	0.674269943	3.22E-13
ORC1	0.672574816	3.88E-13
NETO2	0.671783396	4.23E-13
DEPDC1	0.667912635	6.44E-13
CEP55	0.666856354	7.21E-13
FOXM1	0.666728168	7.31E-13
CCL21	-0.666630276	7.39E-13
ANLN	0.665114529	8.68E-13
KPNA2	0.663190394	1.06E-12
PRR11	0.661162975	1.32E-12
RAD51AP1	0.660009482	1.49E-12
EXO1	0.658274715	1.78E-12
BUB1	0.656559851	2.13E-12
MCM10	0.655964537	2.26E-12
CDCA8	0.655085616	2.47E-12
MELK	0.650163238	4.08E-12
TPX2	0.648992187	4.59E-12
CDC20	0.645462979	6.51E-12
PKMYT1	0.644043346	7.49E-12
UBE2C	0.64339173	7.99E-12
HOXC9	0.642847849	8.43E-12
CDC45	0.640525914	1.06E-11
AURKB	0.640352491	1.07E-11
MAD2L1	0.626883559	3.85E-11
TTK	0.626520302	3.98E-11
NEK2	0.625181801	4.50E-11
KNL1	0.621037878	6.56E-11
FANCB	0.621028082	6.57E-11
KIF23	0.617650486	8.90E-11
CCNE1	0.615384671	1.09E-10
XRCC2	0.612388488	1.42E-10
CCNB2	0.607891867	2.10E-10
KIF18A	0.60766963	2.14E-10
UBE2T	0.605780119	2.52E-10
CENPE	0.602899498	3.22E-10
GRIN2D	0.602241513	3.40E-10
DPF1	0.599447386	4.31E-10
TROAP	0.597523292	5.06E-10
DLGAP5	0.597293989	5.16E-10
CKAP2L	0.595268926	6.10E-10
BUB1B	0.59351203	7.05E-10
ASPM	0.590520003	9.01E-10

CENPF	0.588970946	1.02E-09
SPAG5	0.583323093	1.61E-09
PRC1	0.582283268	1.74E-09
CDCA2	0.581294502	1.89E-09
FANCI	0.580611501	1.99E-09
CDK1	0.577487543	2.54E-09
RRM2	0.575795387	2.90E-09
SKA1	0.574036688	3.32E-09
CHRNA5	0.573969723	3.34E-09
ZIC5	0.573427254	3.48E-09
HOXC8	0.572893976	3.63E-09
PNPLA3	0.572266688	3.80E-09
NXPH4	0.56742585	5.49E-09
NUSAP1	0.567055273	5.65E-09
PHLDB2	0.559933909	9.59E-09
HOXC5	0.555921803	1.29E-08
TMEM158	0.553076781	1.58E-08
PBK	0.54853927	2.18E-08
NDC80	0.534290781	5.84E-08
TOP2A	0.53110395	7.23E-08
NPNT	0.521939478	1.32E-07
STEAP4	-0.51821924	1.68E-07
RPL39L	0.509591407	2.90E-07
TICRR	0.495440497	6.88E-07
ZIC2	0.489523957	9.76E-07
HOXC10	0.489194407	9.95E-07
SLCO1A2	0.481571208	1.55E-06
DUSP9	0.478233158	1.87E-06
GPNMB	0.471407278	2.74E-06
ODC1	0.445223814	1.10E-05
KRTAP4-1	0.443157552	1.22E-05
IRX4	0.439505895	1.46E-05
C6orf223	0.437653858	1.61E-05
IGF2BP3	0.434852814	1.85E-05
TLX2	0.417481007	4.24E-05
MAGEA6	0.404354841	7.73E-05
SNCB	0.395681557	0.000113
FEZF1	0.390310613	0.000143
TLX3	0.3821646	0.000202
OR2B6	0.367801927	0.000362
OR6C2	0.353787633	0.000624
MAGEB17	0.350054374	0.000719
MAGEA2B	0.314895847	0.002504

Note: “geneSymbol” is gene symbol; “GS.LSCC” is the gene significance of laryngeal squamous cell carcinoma; “p.GS.LSCC” is the *p* value of GS.LSCC.

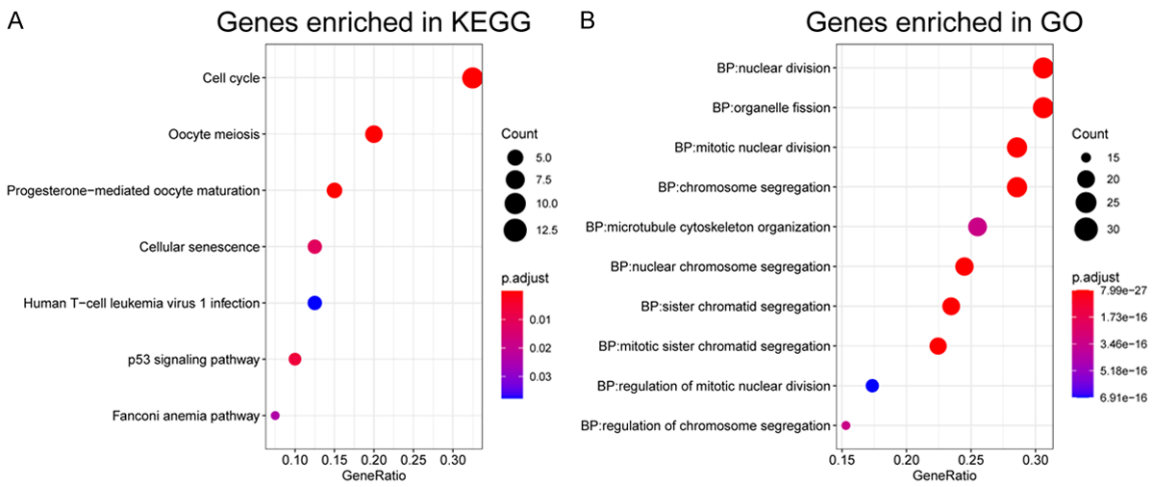

Figure S5. Enriched pathways for genes in module yellow from WGCNA in GSE142083. (A) is from KEGG (Kyoto Encyclopedia of Genes and Genomes) and (B) from GO (Gene Ontology).

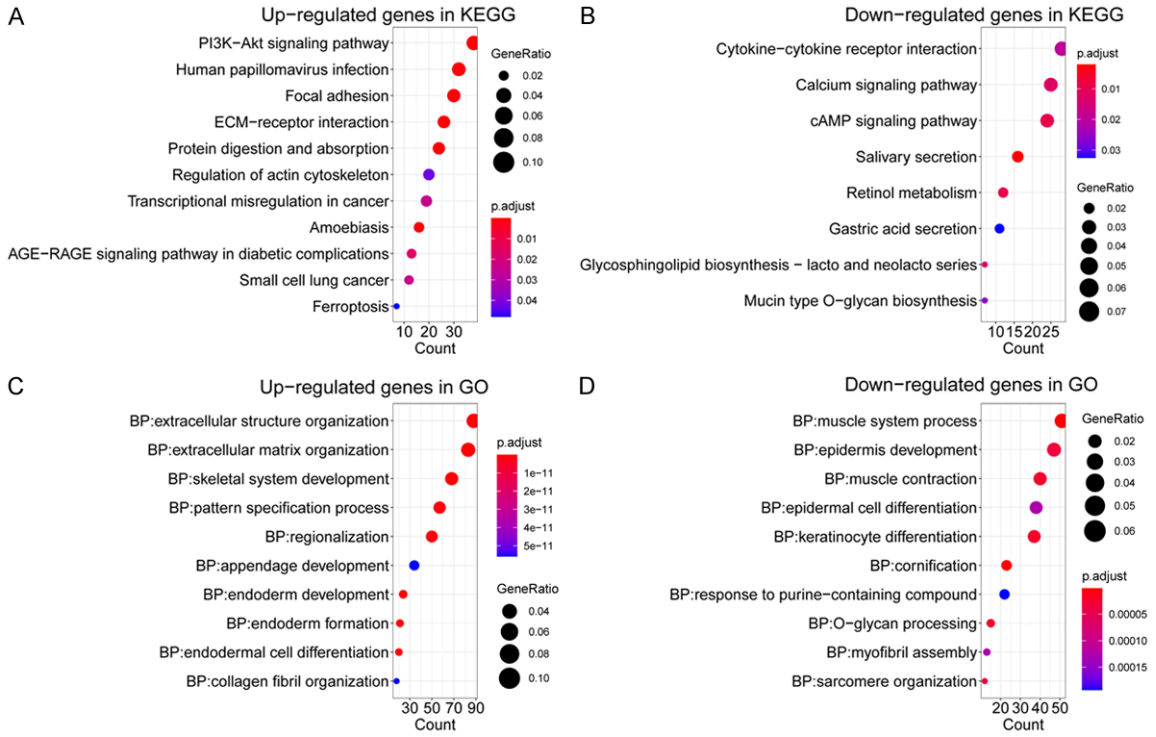

Figure S6. Enriched pathways for TCGA datasets. (A, B) are from KEGG where (A) represents upregulated genes and (B) downregulated genes. (C, D) from GO where (C) represents upregulated genes and (D) downregulated genes.